

# A simple proof of almost percolation on $G(n, p)$ <sup>1</sup>

Mihyun Kang<sup>2</sup>

Tamás Makai<sup>2</sup>

August 3, 2016

## Abstract

We consider bootstrap percolation on the binomial random graph  $G(n, p)$  with infection threshold  $r \in \mathbb{N}$ , an infection process which starts from a set of initially infected vertices and in each step every vertex with at least  $r$  infected neighbours becomes infected. We improve the results of Janson, Łuczak, Turova, and Valier (2012) by strengthening the probability bounds on the number of infected vertices at the end of the process, using simple arguments based on martingales and giant components.

## 1 Introduction

Bootstrap percolation on a graph with infection threshold  $r \in \mathbb{N}$  is a deterministic infection process which evolves in rounds. In each round every vertex has exactly one of two possible states: it is either infected or uninfected. We denote the set of initially infected vertices by  $A(0)$ . In each round of the process every uninfected vertex  $v$  becomes infected if it has at least  $r$  infected neighbours, otherwise it remains uninfected. Once a vertex has become infected, it remains infected forever. The final infected set is denoted by  $A_f$ .

Bootstrap percolation was introduced by Chalupa, Leath, and Reich [15] in the context of magnetic disordered systems. Since then bootstrap percolation processes and extensions have been used to describe several complex phenomena: from neuronal activity [4, 16] to the dynamics of the Ising model at zero temperature [20].

The dependence of the final number of infected vertices on the set of initially infected vertices have been studied for a variety graphs. These include various deterministic graphs such as trees [10, 19], grids [2, 8, 9, 14, 21], and hypercubes [7]. Several random graph models have also been considered including the binomial random graph [18, 22, 26], random graphs with a given degree sequence [3], random regular graphs [11], the Chung-Lu model [5, 6], preferential attachment graphs [1], and random geometric graphs on the hyperbolic plane [13].

This paper focuses on bootstrap percolation on the binomial random graph  $G(n, p)$ , a graph with vertex set  $[n] := \{1, \dots, n\}$  in which every edge appears independently with probability  $p = p(n)$ . Extending the results of Vallier [26], Janson, Łuczak, Turova, and Vallier [22] analysed bootstrap percolation on  $G(n, p)$ , where the set of initially infected vertices  $A(0)$  is chosen uniformly at random from the vertex sets of size  $a \in [n]$ . For  $r \geq 2$  and  $p$  satisfying both  $p = \omega(n^{-1})$  and  $p = o(n^{-1/r})$ , they showed, among other results, that with probability tending to one as  $n \rightarrow \infty$  either only a few additional vertices are infected or almost every vertex becomes infected, depending on the number of initially infected vertices. In addition they determined the probability of both of these events up to an additive term tending to zero as  $n \rightarrow \infty$ .

They also considered the complementary regime of  $p$ . When  $p = O(n^{-1})$ , with probability tending to one as  $n \rightarrow \infty$ , due to the large number of small components, the only way infection can spread to almost every vertex is if almost every vertex was infected initially. In the  $p = \Theta(n^{-1/r})$  case, with probability tending to one as  $n \rightarrow \infty$  every vertex will become infected if the initial set of infected vertices tends to

<sup>1</sup>An earlier version of this paper appeared in the proceedings of the 27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2016).

<sup>2</sup>Graz University of Technology, Institute of Discrete Mathematics, Steyrergasse 30, Graz, Austria. {kang, makai}@math.tugraz.at. Supported by Austrian Science Fund (FWF): P26826.

infinity. In addition when  $p = \omega(n^{-1/r})$  already  $r$  initially infected vertices spread the infection to every vertex in the graph with probability tending to one as  $n \rightarrow \infty$ . Therefore the most interesting range of  $p$  is when  $p = \omega(n^{-1})$  and  $p = o(n^{-1/r})$  and in this paper we will concentrate on this range.

The main contribution of this paper is a simple proof of almost percolation using martingales and the giant component. We introduce a martingale in order to determine the number of infected vertices during the early stages of the process. Furthermore in the supercritical regime we show that the subgraph spanned by the vertices with  $r - 1$  infected neighbours grows large enough to contain a giant component. The infection of just one vertex in this giant component leads to every vertex in the component becoming infected and we show that this in fact happens with exponentially high probability.

A weaker version of the martingale argument was already used in [22]. However we are unaware of any other proof relying on the appearance of the giant component, making our proof simple and original.

**Main Results.** Throughout the paper we assume that  $r \geq 2$  and that both  $p = \omega(n^{-1})$  and  $p = o(n^{-1/r})$  hold. Moreover, any unspecified limits and asymptotics will be as  $n \rightarrow \infty$ . Throughout the paper we assume that  $n$  is sufficiently large. Let

$$\delta := \max \left\{ (np^r)^{1/(2(r-1))}, (np)^{-1/(4(r-1))} \right\}$$

and note that due to our conditions on  $p$  we have  $\delta = o(1)$ . Set

$$t_0 := \left( (1 + \delta) \frac{(r-1)!}{np^r} \right)^{1/(r-1)}.$$

Let  $\hat{\pi}(t) := \mathbb{P}[\text{Bin}(t, p) \geq r]$  and define

$$a_c := -\min_{t \leq t_0} \frac{n\hat{\pi}(t) - t}{1 - \hat{\pi}(t)}$$

(which is defined essentially as  $a_c^*$  in [22], except we replace  $\text{Po}(tp)$  with  $\text{Bin}(t, p)$ ). In addition denote by  $t_c$  the smallest value  $t$  where this minimum is reached. Similarly to Lemma 9.5 in [22] one can show that

$$t_c = (1 + o(1)) \left( \frac{(r-1)!}{np^r} \right)^{1/(r-1)}$$

and

$$a_c = (1 + o(1)) \left( 1 - \frac{1}{r} \right) t_c.$$

We begin with the size of the final infected set in the subcritical case, in which the number of the initially infected vertices is smaller than the critical value  $a_c$ .

**Theorem 1.** *Let  $\alpha := \alpha(n)$  be any function satisfying the conditions  $\alpha = \omega(\sqrt{a_c})$  and  $\alpha \leq a_c - r$ . If  $|A(0)| = a_c - \alpha$ , then with probability at least*

$$1 - \exp \left( -(1 + o(1)) \frac{r\alpha^2}{2(t_0 + r\alpha/3)} \right)$$

*we have  $|A_f| < t_c$ .*

Following is our main result, which says that in the supercritical case, i.e. when the number of initially infected vertices is larger than  $a_c$ , almost every vertex becomes infected.

**Theorem 2.** *Let  $\alpha := \alpha(n)$  be any function satisfying the conditions  $\alpha = \omega(\sqrt{a_c})$  and  $\alpha \leq t_0 - a_c$ . If  $|A(0)| = a_c + \alpha$ , then with probability at least*

$$1 - \exp \left( -(1 + o(1)) \frac{r\alpha^2}{8(t_0 + r\alpha/3)} \right) - \exp \left( -(1 + o(1)) \frac{(r-1)\alpha^2}{8(t_0 + (r-1)\alpha/2)} \right)$$

*we have  $|A_f| = (1 + o(1))n$ .*

**Proof Technique.** When the number of infected vertices is small (at most  $t_0$ ), we need to perform all calculations very carefully, because a small difference in the number of infected vertices can change the outcome of the process significantly. In order to achieve this we introduce a martingale to show that the number of infected vertices is concentrated around its expectation with *exponentially* high probability. The martingale is similar to the one used in [22], however the maximal one step difference in our martingale is significantly smaller and thus provides tighter concentration (Lemma 7).

In the subcritical regime, the expected number of infected vertices is less than  $t_c \leq t_0$  and therefore the martingale argument alone implies the result (Section 4).

In the supercritical case, the martingale argument is not sufficient as the probability that the martingale is sufficiently concentrated decreases significantly once  $t_0$  vertices have been infected. However it still ensures the number of infected vertices will reach  $t_0$  with exponentially high probability. In fact, at least  $t_0 + \alpha/2$  vertices become infected (Lemma 8). Now take a subset of the infected vertices of size  $t_0 + \alpha/4$  and consider the vertices with at least  $r - 1$  neighbours in that set. The size of this set is roughly  $(1 + \varepsilon)p^{-1}$  for some positive sequence  $\varepsilon := \varepsilon(n, \alpha, \delta)$  (Lemma 9) and the subgraph spanned by these vertices is a binomial random graph,  $G((1 + \varepsilon)p^{-1}, p)$ . A closer examination of  $\varepsilon$  reveals that with respect to the appearance of the giant component, we are in the weakly supercritical regime and a recent result of Bollobás and Riordan [12] (Theorem 5) implies that this random graph has a giant component of size  $2\varepsilon p^{-1}/(1 + 2\varepsilon)$  with exponentially high probability (Lemma 10). Should any vertex in the giant component have an additional infected neighbour, then every vertex in the giant component will become infected eventually. We show that this happens with exponentially high probability. This implies a significant increase in the number of infected vertices and at this point we already have  $\omega(t_0)$  infected vertices. After this, the process ends in two steps, which can be shown by two simple applications of the Chernoff bound (Lemmas 11 and 12).

## 2 Preliminaries

We will use the following form of the Chernoff bound.

**Theorem 3.** [17] *Let  $X \sim \text{Bin}(n, p)$ , i.e. a binomial random variable with parameters  $n$  and  $p$ . Then for any  $\lambda > 0$*

$$\mathbb{P}[X - \mathbb{E}[X] \leq -\lambda] \leq \exp\left(-\frac{\lambda^2}{2\mathbb{E}[X]}\right)$$

and

$$\mathbb{P}[X - \mathbb{E}[X] \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2(\mathbb{E}[X] + \lambda/3)}\right).$$

Let  $M_0, \dots, M_k$  be a sequence of random variables. For each  $0 \leq i \leq k$  denote by  $\mathcal{F}(i)$  the  $\sigma$ -algebra generated by the random variables  $M_0, \dots, M_i$ , i.e.,  $\mathcal{F}(i) = \sigma(\{M_j\}_{0 \leq j \leq i})$ , and we call the sequence  $\{\mathcal{F}(i)\}_{0 \leq i \leq k}$  the natural filtration of  $M_0, \dots, M_k$ .

The following martingale concentration bound which is a slightly generalised form of a result due to Chung and Lu [17] (a strengthened version of McDiarmid [24]) will prove to be vital. For completeness we provide a proof in the appendix.

**Theorem 4.** *For  $m_0 \in \mathbb{R}$  let  $M_0 = m_0, M_1, \dots, M_k$  be a martingale with respect to its natural filtration  $\{\mathcal{F}(i) = \sigma(\{M_j\}_{0 \leq j \leq i})\}_{0 \leq i \leq k}$  whose conditional variance and differences satisfy the following conditions for each  $1 \leq i \leq k$ .*

(i)  $|M_i - M_{i-1}| \leq m$  for some  $m \in \mathbb{R}_+$ ;

(ii)  $\text{Var}[M_i | \mathcal{F}(i-1)] = \mathbb{E}[(M_i - \mathbb{E}[M_i | \mathcal{F}(i-1)])^2 | \mathcal{F}(i-1)] \leq \sigma_i^2$  for some  $\sigma_i \in \mathbb{R}$ .

Then for any  $\lambda > 0$ , we have

$$\mathbb{P}[M_k - m_0 \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2\left(\sum_{i=1}^k \sigma_i^2 + m\lambda/3\right)}\right).$$

We will also need the following result on the appearance of a giant component in  $G(n, p)$  by Bollobás and Riordan, which follows from Theorem 3 in [12].

**Theorem 5.** Let  $\varepsilon = \varepsilon(n)$  satisfy  $\varepsilon = O(1)$  and  $\varepsilon^3 n = \omega(1)$ . Denote by  $\rho \in (0, 1)$  the unique positive solution of  $1 - \rho = \exp(-(1 + \varepsilon)\rho)$ . For any  $0 \leq \gamma < \rho$  with probability  $1 - \exp(-\Omega(\gamma^2 \varepsilon n))$  the binomial random graph  $G(n, (1 + \varepsilon)/n)$  has a component of size at least  $\rho n - \gamma n$ .

### 3 Setup: Martingale

In order to analyse the bootstrap percolation on  $G(n, p)$  we will use the following reformulation similar to Scalia-Tomba [22, 25]. Roughly speaking we examine the infected vertices one by one and determine the vertices which have at least  $r$  neighbours in the set of previously examined vertices. The set of examined vertices until step  $t$  is denoted by  $Z(t)$  and the set of infected vertices by  $A(t)$ .

Formally let  $A(0)$  be the set of initially infected vertices chosen uniformly at random from the vertex sets of size  $a \in [n]$  and without loss of generality we may assume that  $A(0) = \{1, \dots, a\}$ . Set  $Z(0) = \emptyset$ . For each step  $t \in \mathbb{N}$ , if  $A(t-1) \setminus Z(t-1) \neq \emptyset$ , then let  $U(t) = \{u(t)\}$ , where  $u(t)$  is the smallest vertex in  $A(t-1) \setminus Z(t-1)$ , otherwise set  $U(t) = \emptyset$ . Set  $Z(t) := Z(t-1) \cup U(t)$ . Now for  $t \geq 0$  and  $i \in [n-a] := \{1, \dots, n-a\}$  let  $X(t, i)$  be the indicator random variable for the event that the vertex  $a+i$  has at least  $r$  neighbours in  $Z(t)$  and set

$$A(t) := A(0) \cup \{a+i : X(t, i) = 1, i \in [n-a]\}.$$

The process stops when  $t = n$ .

Clearly  $Z(t) \subset A(t)$ . Let  $T$  denote the smallest value of  $t$  such that  $A(t) = Z(t)$ . Note that  $t \leq T$  implies that  $|Z(t)| = t$  and thus  $T$  is also the smallest  $t$ , in fact the only  $t$  satisfying  $|A(t)| = t$ . Since  $|A(t)| \leq n$  for every  $t \in \{0, 1, \dots, n\}$ , we have that  $T \leq n$ . Note further that  $A(T) = A_f$ .

In [22] a martingale was introduced to analyse this process. Note that in any step of the process at most  $n - |A(t)|$  vertices may become infected (which in the early stages of the process is  $(1 + o(1))n$ ) and martingale concentration inequalities depend on the maximal one step difference. Therefore only very weak concentration can be shown for this martingale.

In order to achieve a better control on the maximal one step difference in a martingale, we refine the process by dividing every step into rounds, in such a way that in each round *exactly one vertex*  $v \in [n] \setminus A(0)$  is examined (regardless whether it was examined in earlier rounds or not). Thus each step  $t \leq n$  consists of  $n - a$  rounds and round  $i$  of step  $t$  is denoted by  $(t, i)$ . We denote the round following  $(t, i)$  by  $(t, i) + 1$  and the round preceding  $(t, i)$  by  $(t, i) - 1$ . Also the ordering of the rounds is given by the lexicographical order, i.e.  $(\tau, \iota) < (t, i)$  if either  $\tau < t$  or  $\tau = t$  and  $\iota < i$ .

In round  $i$  of step  $t$  we examine if the vertex  $a+i$  has at least  $r$  neighbours in  $Z(t)$  and if it does we add it to the set of infected vertices. Formally for each  $(t, i) \geq (1, 1)$  we let

$$A(t, i) := A(0) \cup \{a+j : X(t, j) = 1, j = 1, \dots, i\} \cup \{a+j : X(t-1, j) = 1, j = i+1, \dots, n-a\}. \quad (1)$$

Clearly we have  $A(t) = A(t, n-a)$ . For consistency define  $A(0, n-a) := A(0)$ .

Define a function  $\pi : \mathbb{N} \rightarrow [0, 1]$  by

$$\pi(t) := \begin{cases} \mathbb{P}[\text{Bin}(t, p) \geq r], & \text{for } t \leq T; \\ \mathbb{P}[\text{Bin}(T, p) \geq r], & \text{for } t > T. \end{cases}$$

Note that  $\pi(t)$  is a random variable.

For  $(t, i) \geq (0, n-a)$ , define the random variable

$$M(t, i) := \sum_{j=1}^i \frac{X(t, j) - \pi(t)}{1 - \pi(t)} + \sum_{j=i+1}^{n-a} \frac{X(t-1, j) - \pi(t-1)}{1 - \pi(t-1)}. \quad (2)$$

Since  $\pi(t) < 1$  for every  $t \in \mathbb{N}$ , the random variable  $M(t, i)$  is well defined. Tedious arguments show that the sequence of random variables  $M(0, n-a), \dots, M(n, n-a)$  forms a martingale (see the appendix for the proof).

**Lemma 6.** For each  $(0, n-a) \leq (t, i) \leq (n, n-a)$  we denote by  $\mathcal{F}(t, i)$  the  $\sigma$ -algebra generated by  $M(0, n-a), \dots, M(t, i)$ . The sequence of  $M(0, n-a), \dots, M(n, n-a)$  forms a martingale with respect to the filtration  $\{\mathcal{F}(0, n-a), \dots, \mathcal{F}(n, n-a)\}$ .

Next we show that  $M(t, i)$  is concentrated around its expectation.

**Lemma 7.** Let  $t \in \{0, \dots, n\}$  and  $\lambda \in \mathbb{R}_+$  be given. Then we have

$$\mathbb{P} \left[ \bigwedge_{(0, n-a) \leq (\tau, i) \leq (t, n-a)} (M(\tau, i) > -\lambda) \right] \geq 1 - \exp \left( -\frac{\lambda^2(1 - \hat{\pi}(t))^3}{2(n\hat{\pi}(t) + \lambda/3)} \right)$$

and

$$\mathbb{P} \left[ \bigwedge_{(0, n-a) \leq (\tau, i) \leq (t, n-a)} (M(\tau, i) < \lambda) \right] \geq 1 - \exp \left( -\frac{\lambda^2(1 - \hat{\pi}(t))^3}{2(n\hat{\pi}(t) + \lambda/3)} \right).$$

*Proof.* We only need to consider the bound on the probability that  $M(\tau, i) < \lambda$  for each  $(0, n-a) \leq (\tau, i) \leq (t, n-a)$ . The other case follows simply from the fact that if the sequence of random variables  $M(0, n-a), \dots, M(t, n-a)$  forms a martingale with respect to a filtration, then  $-M(0, n-a), \dots, -M(t, n-a)$  is also a martingale with respect to the same filtration and they both have the same conditional variance and maximal difference. In order to show that the bounds hold for each step, we introduce the following martingale

$$\hat{M}(\tau, i) = \begin{cases} M(\tau, i) & \text{if } \hat{M}((\tau, i) - 1) < \lambda, \\ \hat{M}((\tau, i) - 1) & \text{otherwise.} \end{cases}$$

Similarly to  $M(\tau, i)$  we denote by  $\hat{\mathcal{F}}(\tau, i)$  the  $\sigma$ -algebra generated by  $\hat{M}(0, n-a), \dots, \hat{M}(\tau, i)$ . Note that if there exists a round  $(\tau, i)$  such that  $M(\tau, i) \geq \lambda$ , then we have  $\hat{M}(\tau', i') \geq \lambda$  for every  $(\tau', i') \geq (\tau, i)$ . Therefore  $\hat{M}(t, n-a) < \lambda$  implies that for every  $(0, n-a) \leq (\tau, i) \leq (t, n-a)$  we have  $M(\tau, i) < \lambda$ . Thus it suffices to show that

$$\mathbb{P} [\hat{M}(t, n-a) \geq \lambda] \leq \exp \left( -\frac{\lambda^2(1 - \hat{\pi}(t))^3}{2(n\hat{\pi}(t) + \lambda/3)} \right). \quad (3)$$

Our aim is to apply Theorem 4 to  $\hat{M}(\tau, i)$ . For this we need an upper bound on the maximal one step difference and conditional variance.

First we give an upper bound on the one step difference. We can easily compute

$$M(t, i) - M((t, i) - 1) \stackrel{(2)}{=} \frac{X(t, i) - \pi(t)}{1 - \pi(t)} - \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)}. \quad (4)$$

By (4) and  $\pi(\tau) \leq \hat{\pi}(\tau)$ , we have for  $\tau \leq t$

$$|\hat{M}(\tau, i) - \hat{M}((\tau, i) - 1)| \leq \max \left\{ \frac{\hat{\pi}(\tau)}{1 - \hat{\pi}(\tau)} - \frac{\hat{\pi}(\tau-1)}{1 - \hat{\pi}(\tau-1)}, 1 + \frac{\hat{\pi}(\tau-1)}{1 - \hat{\pi}(\tau-1)} \right\}.$$

Since  $\hat{\pi}(\tau) = \mathbb{P}[\text{Bin}(\tau, p) \geq r]$  we have

$$\begin{aligned} \hat{\pi}(\tau) - \hat{\pi}(\tau-1) &= \binom{\tau-1}{r-1} p^r (1-p)^{\tau-r} \leq \binom{\tau}{r-1} p^r (1-p)^{\tau-r} \\ &\stackrel{p=O(1)}{\leq} \binom{\tau}{r-1} p^{r-1} (1-p)^{\tau-r+1} \leq \sum_{i=0}^{r-1} \binom{\tau}{i} p^i (1-p)^{\tau-i} = 1 - \hat{\pi}(\tau) \end{aligned}$$

and consequently

$$\frac{\hat{\pi}(\tau)}{1 - \hat{\pi}(\tau)} - \frac{\hat{\pi}(\tau-1)}{1 - \hat{\pi}(\tau-1)} \leq 1 + \frac{\hat{\pi}(\tau-1)}{1 - \hat{\pi}(\tau-1)}.$$

Therefore we obtain

$$|\hat{M}(\tau, i) - \hat{M}((\tau, i) - 1)| \leq 1 + \frac{\hat{\pi}(\tau-1)}{1 - \hat{\pi}(\tau-1)} = \frac{1}{1 - \hat{\pi}(\tau-1)} \stackrel{\tau \leq t}{\leq} \frac{1}{1 - \hat{\pi}(t)}. \quad (5)$$

Next we examine the conditional variance. Conditional on  $\hat{M}((\tau, i) - 1) \geq \lambda$ , the random variable  $\text{Var}[\hat{M}(\tau, i) | \hat{\mathcal{F}}((\tau, i) - 1)] = 0$ , on the other hand conditional on  $\hat{M}((\tau, i) - 1) < \lambda$  we have

$$\text{Var}[\hat{M}(\tau, i) | \hat{\mathcal{F}}((\tau, i) - 1), \hat{M}((\tau, i) - 1) < \lambda] = \text{Var}[M(\tau, i) | \mathcal{F}((\tau, i) - 1)].$$

Note that

$$\text{Var}[M(\tau, i)|\mathcal{F}((\tau, i) - 1)] = \text{Var}\left[\frac{X(\tau, i)}{1 - \pi(t)} \middle| \mathcal{F}((\tau, i) - 1)\right] \leq \frac{1}{(1 - \hat{\pi}(t))^2} \text{Var}[X(\tau, i)|\mathcal{F}((\tau, i) - 1)].$$

We shall show that

$$\text{Var}[X(\tau, i)|\mathcal{F}((\tau, i) - 1)] \leq \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau)}. \quad (6)$$

Recall that  $X(\tau - 1, i) = 1$  implies  $X(\tau, i) = 1$  and that  $\tau > T$  implies  $X(\tau, i) = X(\tau - 1, i)$ . In both of these cases we have

$$\begin{aligned} \text{Var}[X(\tau, i)|\mathcal{F}((\tau, i) - 1), X(\tau - 1, i) = 1] &= \text{Var}[X(\tau, i)|\mathcal{F}((\tau, i) - 1), \tau > T] = 0 \\ &\stackrel{\hat{\pi}(\tau) \geq \hat{\pi}(\tau - 1)}{\leq} \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau)}. \end{aligned}$$

Now assume  $\tau \leq T$  and  $X(\tau - 1, i) = 0$ . Since  $X(\tau, i)$  is an indicator random variable, we have

$$\begin{aligned} \text{Var}[X(\tau, i)|\mathcal{F}((\tau, i) - 1), \tau \leq T, X(\tau - 1, i) = 0] &\leq \mathbb{E}[X(\tau, i)|\mathcal{F}((\tau, i) - 1), \tau \leq T, X(\tau - 1, i) = 0] \\ &\stackrel{(24)}{=} \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau - 1)} \leq \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{1 - \hat{\pi}(\tau)} \end{aligned}$$

and thus (6) holds. Therefore we have

$$\begin{aligned} \sum_{\tau=1}^t \sum_{i=1}^{n-a} \text{Var}[M(\tau, i)|\mathcal{F}((\tau, i) - 1)] &\stackrel{(6)}{\leq} \sum_{\tau=1}^t \sum_{i=1}^{n-a} \frac{\hat{\pi}(\tau) - \hat{\pi}(\tau - 1)}{(1 - \hat{\pi}(\tau))^3} \\ &\leq \sum_{\tau=1}^t \frac{n(\hat{\pi}(\tau) - \hat{\pi}(\tau - 1))}{(1 - \hat{\pi}(\tau))^3} \leq \frac{n\hat{\pi}(t)}{(1 - \hat{\pi}(t))^3}. \end{aligned} \quad (7)$$

Note that  $M(0, n - a) = 0$ . Thus by Theorem 4 with (5) and (7) we have

$$\mathbb{P}[\hat{M}(t, n - a) \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2} \left(\frac{n\hat{\pi}(t)}{(1 - \hat{\pi}(t))^3} + \frac{\lambda}{3(1 - \hat{\pi}(t))}\right)^{-1}\right)$$

implying (3). This completes the proof.  $\square$

The previous lemma allows us to analyse the process in the first  $t_0$  steps. This will be used in the proofs of Theorems 1 and 2.

## 4 Subcritical case: Proof of Theorem 1

We want to investigate the number of infected vertices at time  $t_c$ . By the definition of  $a_c$  and  $t_c$ , we have

$$a_c = -\min_{t \leq t_0} \frac{n\hat{\pi}(t) - t}{1 - \hat{\pi}(t)} = \frac{t_c - n\hat{\pi}(t_c)}{1 - \hat{\pi}(t_c)}. \quad (8)$$

Using (1) and (2) we can express the number of infected vertices until step  $t$  by

$$|A(t)| \stackrel{(1)}{=} a + \sum_{i=1}^{n-a} X(t, i) \stackrel{(2)}{=} a + M(t, n - a)(1 - \pi(t)) + (n - a)\pi(t) \quad (9)$$

and in particular we have for step  $t_c$

$$|A(t_c)| = a + (1 - \pi(t_c))M(t_c, n - a) + (n - a)\pi(t_c).$$

Since  $\pi(t) \leq \hat{\pi}(t)$  and  $a = a_c - \alpha$ , we obtain

$$\begin{aligned} |A(t_c)| &\leq a + M(t_c, n - a) + (n - a)\hat{\pi}(t_c) \\ &= (a_c - \alpha)(1 - \hat{\pi}(t_c)) + n\hat{\pi}(t_c) + M(t_c, n - a) \\ &\stackrel{(8)}{=} t_c - n\hat{\pi}(t_c) + n\hat{\pi}(t_c) - \alpha(1 - \hat{\pi}(t_c)) + M(t_c, n - a) \\ &= t_c - \alpha(1 - \hat{\pi}(t_c)) + M(t_c, n - a). \end{aligned} \quad (10)$$

Since  $t_0 = (1 + o(1))((r-1)!/(np^r))^{1/(r-1)}$ , we have

$$t_0 p = O\left(\left(\frac{1}{np^r}\right)^{1/(r-1)} p\right) = O\left(\left(\frac{1}{np}\right)^{1/(r-1)}\right) \stackrel{np=\omega(1)}{=} o(1). \quad (11)$$

Furthermore,

$$\begin{aligned} \hat{\pi}(t_0) &= \mathbb{P}[\text{Bin}(t_0, p) \geq r] = \sum_{j=r}^{t_0} \binom{t_0}{j} p^j (1-p)^{t_0-j} \stackrel{(11)}{=} (1+o(1)) \frac{t_0^r p^r}{r!} \\ &= (1+o(1)) \frac{t_0^{r-1} p^r}{r!} t_0 = (1+o(1)) \frac{1}{r} \frac{t_0}{n} \stackrel{np=\omega(1)}{=} o(t_0 p) \stackrel{(11)}{=} o(1). \end{aligned} \quad (12)$$

Applying Lemma 7 with  $\lambda = (1 - \sqrt{\hat{\pi}(t_c)})\alpha$ , we have

$$M(t_c, n-a) < (1 - \sqrt{\hat{\pi}(t_c)})\alpha \quad (13)$$

with probability at least

$$\begin{aligned} 1 - \exp\left(- (1+o(1)) \frac{\alpha^2}{2(n\hat{\pi}(t_c) + \alpha/3)}\right) &\stackrel{\hat{\pi}(t_c) \leq \hat{\pi}(t_0)}{\geq} 1 - \exp\left(- (1+o(1)) \frac{\alpha^2}{2(n\hat{\pi}(t_0) + \alpha/3)}\right) \\ &\stackrel{(12)}{=} 1 - \exp\left(- (1+o(1)) \frac{r\alpha^2}{2(t_0 + r\alpha/3)}\right). \end{aligned}$$

Furthermore, (10) and (13) imply that with probability at least  $1 - \exp\left(- (1+o(1)) \frac{r\alpha^2}{2(t_0 + r\alpha/3)}\right)$  we have

$$|A(t_c)| \leq t_c - (1 - \hat{\pi}(t_c))\alpha + (1 - \sqrt{\hat{\pi}(t_c)})\alpha = t_c + (\hat{\pi}(t_c) - \sqrt{\hat{\pi}(t_c)})\alpha \stackrel{\hat{\pi}(t_c) < 1}{<} t_c$$

and therefore  $|A_f| = T < t_c$ , as desired.

## 5 Supercritical case: Proof of Theorem 2

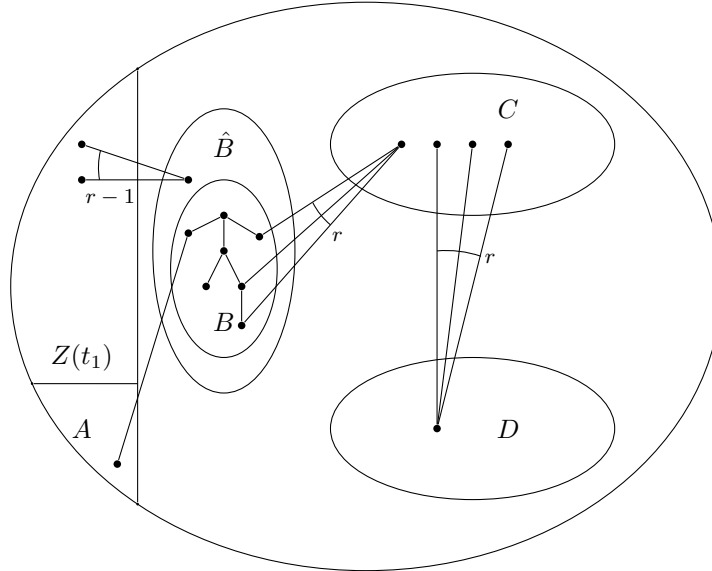


Figure 1: Spread of the infection

For the remainder of the paper fix an  $\alpha$  satisfying the conditions of Theorem 2, i.e.,  $\alpha = \omega(\sqrt{a_c})$  and  $\alpha \leq t_0 - a_c$ . Define

$$t_1 := t_0 + \alpha/4.$$

The proof of Theorem 2 consists of several steps (see Figure 1). First we show that the process runs for at least  $t_1$  steps and in fact at this stage there exists a ‘large’ set of vertices  $A \subset A(t_1) \setminus Z(t_1)$  (Lemma 8). Next we consider the set  $\hat{B}$  of vertices which have at least  $r - 1$  neighbours in  $Z(t_1)$  and prove a lower bound on the number of these vertices (Lemma 9). It turns out that the size of  $\hat{B}$  is large enough for it to contain a giant component  $B$  (Lemma 10). If there exists a vertex in  $A$  which is connected to a vertex in  $B$ , then every vertex in  $B$  will become infected. Finally we examine the set  $C$  of vertices which have at least  $r$  neighbours in  $B$  (Lemma 11) and the set  $D$  of vertices which have at least  $r$  neighbours in  $C$  (Lemma 12). We complete the proof by showing  $|D| = (1 + o(1))n$ , because  $D \subset A_f$ .

**Lemma 8.** *If  $|A(0)| = a_c + \alpha$ , then with probability at least*

$$1 - \exp\left(- (1 + o(1)) \frac{r\alpha^2}{8(t_0 + r\alpha/3)}\right)$$

*we have  $T > t_1$  and  $|A(t_1)| \geq t_1 + (1 - 2\hat{\pi}(t_0))\alpha/4$ .*

*Proof.* From (12) and Lemma 7 with  $\lambda = \alpha/2$  we have

$$M(t, i) > -\alpha/2, \quad \forall (0, n - a) \leq (t, i) \leq (t_0, n - a) \quad (14)$$

with probability at least

$$1 - \exp\left(- (1 + o(1)) \frac{r\alpha^2}{8(t_0 + r\alpha/3)}\right).$$

Conditional on this event we shall show that  $T > t_0$  and  $|A(t_0)| \geq t_0 + (1 - \hat{\pi}(t_0))\alpha/2$ . By the definition of  $a_c$ , for every  $t \leq t_0$  we have

$$a_c \geq \frac{t - n\hat{\pi}(t)}{1 - \hat{\pi}(t)}. \quad (15)$$

We will indeed show by induction that for every  $0 \leq t \leq t_0$  we have that  $|A(t)| \geq t + (1 - \hat{\pi}(t))\alpha/2$ . Note that this implies that  $|A(t)| > t$  for every  $0 \leq t \leq t_0$  and thus  $T > t_0$ . Clearly  $|A(0)| = a_c + \alpha \geq \alpha/2$ . Now assume that for some  $0 \leq t \leq t_0$  we have for every  $\tau < t$  that  $|A(\tau)| \geq \tau + (1 - \hat{\pi}(\tau))\alpha/2$ . This implies that  $T > t - 1$  and thus  $\pi(t) = \hat{\pi}(t)$ . Therefore we obtain

$$\begin{aligned} |A(t)| &\stackrel{(9)}{=} a + (1 - \pi(t))M(t, n - a) + (n - a)\pi(t) \\ &\stackrel{(14)}{>} (1 - \hat{\pi}(t))(a_c + \alpha) + n\hat{\pi}(t) - (1 - \hat{\pi}(t))\alpha/2 \\ &\stackrel{(15)}{\geq} t + (1 - \hat{\pi}(t))\alpha/2. \end{aligned}$$

For every  $t_1 \geq t \geq t_0$  we have  $|A(t)| \geq |A(t_0)| \geq t_0 + (1 - \hat{\pi}(t_0))\alpha/2$ , and thus

$$|A(t)| \geq t + (t_0 - t_1) + (1 - \hat{\pi}(t_0))\alpha/2 = t + (1 - 2\hat{\pi}(t_0))\alpha/4,$$

which implies the result.  $\square$

Next we shall establish the size of the giant component in the set of vertices which have at least  $r - 1$  neighbours in  $Z(t_1)$ . For this we first need to establish the total number of vertices which have at least  $r - 1$  neighbours in  $Z(t_1)$ .

**Lemma 9.** *Conditional on  $T > t_1$  and  $|A(t_1)| \geq t_1 + (1 - 2\hat{\pi}(t_0))\alpha/4$ , we have that for any set  $A \subset [n] \setminus Z(t_1)$  satisfying  $|A| = (1 - 2\hat{\pi}(t_0))\alpha/4$ , with probability at least*

$$1 - \exp(-\omega(\alpha^2/t_0))$$

*the set of vertices in  $[n] \setminus (Z(t_1) \cup A)$  with at least  $r - 1$  neighbours in  $Z(t_1)$  is at least*

$$\left(1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0}\right)p^{-1}.$$



*Proof.* Let  $\mathcal{E}$  be the event that the set of vertices in  $[n] \setminus (Z(t_1) \cup A)$  with at least  $r-1$  neighbours in  $Z(t_1)$  is at least  $\left(1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0}\right)p^{-1}$ . Since  $A(0)$  contains the smallest vertices in the graph and in every step of the process the smallest vertex from  $A(t) \setminus Z(t)$  is selected we have that  $A(0) = Z(a) \subset Z(t_1)$ . By definition  $Z(t_1) \subset A(t_1)$ . Therefore the result follows once we show that for every  $Z', A'$  satisfying  $A(0) \subset Z' \subset A'$ ,  $|Z'| = t_1$  and  $|A'| \geq t_1 + (1 - 2\hat{\pi}(t_0))\alpha/4$  the probability that  $\mathcal{E}$  holds conditional on the event  $Z(t_1) = Z'$  and  $A(t_1) = A'$  is at least

$$1 - \exp(-\omega(\alpha^2/t_0)).$$

For the remainder of the proof fix  $Z'$  and  $A'$  which satisfy the previous conditions. Define bootstrap percolation restricted to  $A'$  as follows. Let  $A'(0) = A(0)$  and  $Z'(0) = \emptyset$ . For each step  $t \in \mathbb{N}$ , if  $A'(t-1) \setminus Z'(t-1) \neq \emptyset$ , then let  $U'(t) = \{u'(t)\}$ , where  $u'(t)$  is the smallest vertex in  $A'(t-1) \setminus Z'(t-1)$ , otherwise set  $U'(t) = \emptyset$ . Set  $Z'(t) := Z'(t-1) \cup U'(t)$ . In addition  $A'(t)$  contains the set of vertices in  $A'$  which have at least  $r$  neighbours in  $Z'(t)$  and the set of vertices in  $A'(0)$ .

Let  $\mathcal{G}$  be the event that every vertex outside of  $A'$  has less than  $r$  neighbours in  $Z'$ . We will show that the event  $Z(t_1) = Z' \cap A(t_1) = A'$  is equivalent to the event  $Z'(t_1) = Z' \cap A'(t_1) = A' \cap \mathcal{G}$ . Note that if  $Z(t_1) = Z' \cap A(t_1) = A'$  holds, then any vertex in  $[n] \setminus A'$ , as it is not contained in  $A(t_1) = A'$ , can have at most  $r-1$  neighbours in  $Z' = Z(t_1)$ . Therefore the event  $Z(t_1) = Z' \cap A(t_1) = A'$  is contained in  $\mathcal{G}$ .

Next we show that if  $\mathcal{G}$  holds, then for every  $t \leq t_1$  we have that  $Z(t) = Z'(t)$  and  $A(t) = A'(t)$ . Clearly this holds for  $t = 0$  and now assume that it holds until step  $t-1$ . Recall that  $Z(t) = Z(t-1) \cup U(t)$  and  $Z'(t) = Z'(t-1) \cup U'(t)$ . According to our assumption  $Z(t-1) = Z'(t-1)$ , so we only need to show that  $U(t) = U'(t)$ . Recall that  $U(t)$  and  $U'(t)$  are the smallest vertices in the set  $A(t-1) \setminus Z(t-1)$  and  $A'(t-1) \setminus Z'(t-1)$  respectively. Because  $A(t-1) \setminus Z(t-1) = A'(t-1) \setminus Z'(t-1)$ , we have  $U(t) = U'(t)$ , which implies  $Z(t) = Z'(t)$ . It only remains to show that  $A(t) = A'(t)$ . Since  $Z(t) \subset Z(t_1) = Z'$  and  $\mathcal{G}$  holds, we have that any vertex in  $A(t)$  must be in  $A'$ , thus  $A(t)$  is the set of vertices in  $A'$  with at least  $r$  neighbours in  $Z(t)$  and  $A(0)$ , matching the definition of  $A'(t)$ .

Let  $X_v$  be the indicator random variable that a vertex  $v \in [n] \setminus (A \cup Z')$  has at least  $r-1$  neighbours in  $Z'$  and set  $X := \sum_{v \in [n] \setminus (A \cup Z')} X_v$ . For a vertex  $v \in A'$  we have

$$\mathbb{P}[X_v = 1 | A(t_1) = A', Z(t_1) = Z'] = 1.$$

On the other hand, when  $v \notin A'$ , since the event  $Z(t_1) = Z' \cap A(t_1) = A'$  is equivalent to the event  $Z'(t_1) = Z' \cap A'(t_1) = A' \cap \mathcal{G}$ , we have

$$\mathbb{P}[X_v = 1 | A(t_1) = A', Z(t_1) = Z'] = \mathbb{P}[X_v = 1 | A'(t_1) = A', Z'(t_1) = Z', \mathcal{G}].$$

Note that  $A'(t_1)$  and  $Z'(t_1)$  depend only on edges spanned by  $A'$ , while the events  $X_v = 1$  and  $\mathcal{G}$  depend only on edges with exactly one end in  $A'$ . Therefore the event  $Z'(t_1) = Z' \cap A'(t_1) = A'$  is independent of the event  $X_v = 1 \cap \mathcal{G}$  and the event  $\mathcal{G}$  as well. Therefore

$$\mathbb{P}[X_v = 1 | A'(t_1) = A', Z'(t_1) = Z', \mathcal{G}] = \mathbb{P}[X_v = 1 | \mathcal{G}].$$

Denote by  $d_{Z'}(u)$  the number of neighbours of  $u$  in  $Z'$ . We have

$$\mathcal{G} = \{\forall u \in [n] \setminus A' : d_{Z'}(u) < r\}.$$

Note that the event  $d_{Z'}(u)$  depends only on edges between  $u$  and  $Z'$  and the event  $X_v = 1$  depends only on edges between  $v$  and  $Z'$ . Since these events are independent unless  $v = u$ , we have

$$\mathbb{P}[X_v = 1 | \mathcal{G}] = \mathbb{P}[X_v = 1 | d_{Z'}(v) < r] \geq \mathbb{P}[X_v = 1, d_{Z'}(v) < r].$$

In addition, if  $v$  has at least  $r-1$  neighbours in  $Z'$  and  $d_{Z'}(v) < r$ , then  $v$  has exactly  $r-1$  neighbours in  $Z'$  and thus for  $v \notin A'$  we have

$$\begin{aligned} \mathbb{P}[X_v = 1, d_{Z'}(v) < r] &= \binom{t_1}{r-1} p^{r-1} (1-p)^{t_1-r+1} \\ &= (1 + O(t_1^{-1})) \frac{t_1^{r-1}}{(r-1)!} p^{r-1} (1-p)^{t_1-r+1}. \end{aligned}$$

Since  $t_1 = O(t_0)$ , it follows from (11) that  $(1-p)^{t_1-r+1} = 1 + O(t_0p)$  and thus

$$\begin{aligned}\mathbb{P}[X_v = 1, d_{Z'}(v) < r] &\geq (1 + O(t_1^{-1}) + O(t_0p)) \frac{t_0^{r-1}(1 + (r-1)\alpha/(4t_0))}{(r-1)!} p^{r-1} \\ &\geq (1 + O(t_1^{-1}) + O(t_0p)) \left( \frac{1+\delta}{np} + \frac{(r-1)\alpha}{4npt_0} \right).\end{aligned}$$

Since  $t_1^{-1} = O(t_0^{-1}) = o(\delta)$ ,  $t_0p = o(\delta)$  and  $\alpha = O(t_0)$ , we have

$$\mathbb{P}[X_v = 1 | A(t_1) = A', Z(t_1) = Z'] \geq \frac{1+\delta+o(\delta)}{np} + \frac{(r-1)\alpha}{4npt_0}.$$

Conditional on  $Z'(t_1) = Z'$ ,  $A'(t_1) = A'$  and  $\mathcal{G}$ , the set of random variables  $\{X_v | v \in [n] \setminus A'\}$  are mutually independent. Also conditional on these events we have that  $X_v = 1$  when  $v \in A' \setminus (A \cup Z')$  and thus the set of random variables  $\{X_v | v \in [n] \setminus (Z' \cup A)\}$  are also mutually independent. Therefore  $X$  stochastically dominates the binomial random variable

$$\hat{X} = \text{Bin} \left( n - t_1 - |A|, \frac{1+\delta+o(\delta)}{np} + \frac{(r-1)\alpha}{4npt_0} \right).$$

Since  $(t_1 + |A|)/n = O(t_1/n) = O((np)^{-r/(r-1)}) = o(\delta)$ , we have

$$\mathbb{E}[\hat{X}] = \frac{1+\delta+o(\delta)}{p} + \frac{(r-1)\alpha}{4pt_0}.$$

Recall that  $\alpha \leq t_0$  and thus  $\mathbb{E}[\hat{X}] \leq rp^{-1}$ . This and Theorem 3 imply

$$\mathbb{P} \left[ X - \mathbb{E}(\hat{X}) \leq -\frac{\alpha}{(t_0p)^{2/3}} \right] \leq \exp \left( -\frac{1}{2} \frac{\alpha^2}{(t_0p)^{4/3}} \frac{1}{rp^{-1}} \right) = \exp \left( -\Omega \left( \frac{\alpha^2}{t_0(t_0p)^{1/3}} \right) \right) \stackrel{(11)}{=} \exp \left( -\omega \left( \frac{\alpha^2}{t_0} \right) \right).$$

We only need to show that

$$\mathbb{E}[\hat{X}] - \frac{\alpha}{(t_0p)^{2/3}} = \frac{1+\delta+o(\delta)}{p} + \frac{(r-1)\alpha}{4pt_0} - \frac{\alpha}{(t_0p)^{2/3}} \geq \left( 1 + \frac{3\delta}{4} + \frac{(r-1)\alpha}{4t_0} \right) p^{-1}.$$

The result follows from the fact that

$$\frac{\alpha}{(t_0p)^{2/3}} = \frac{\alpha}{t_0p} (t_0p)^{1/3} = O((t_0p)^{1/3} p^{-1}) = O((np)^{-1/(3(r-1))} p^{-1}) = o(\delta p^{-1}).$$

□

Next we establish the size of a giant component in a set of size  $\left( 1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0} \right) p^{-1}$ .

**Lemma 10.** *With probability  $1 - \exp(-\omega(\alpha^2/t_0))$  the binomial random graph*

$$G \left( \left( 1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0} \right) p^{-1}, p \right)$$

*contains a component of size at least*

$$\left( \frac{\delta}{4} + \frac{(r-1)\alpha}{2t_0 + (r-1)\alpha} \right) p^{-1}.$$

*Proof.* Recall that  $\delta \geq (np)^{-1/(4(r-1))}$  and thus

$$\delta^3 p^{-1} \geq \left( \frac{1}{np} \right)^{3/(4(r-1))} p^{-1} = \left( \frac{1}{np^r} \right)^{3/(4(r-1))} p^{-1/4} = \Omega \left( \left( t_0^{3/4} p^{-1/4} \right) \right) \stackrel{(11)}{=} \omega(t_0) = \omega(1). \quad (16)$$

Therefore we are in the range where Theorem 5 is applicable. Note that for any  $\varepsilon > 0$  and any  $x < (2\varepsilon)/(1 + 2\varepsilon)$  we have

$$\exp((1 + \varepsilon)x) > \sum_{k=0}^2 \frac{((1 + \varepsilon)x)^k}{k!} > 1 + x + x^2 \frac{1}{1 - x} = \frac{1}{1 - x}.$$

Thus the unique positive solution  $\rho$  of the equation  $1 - \rho = \exp(-\rho(1 + 3\delta/4 + (r - 1)\alpha/(4t_0)))$  satisfies

$$\rho > \frac{6\delta/4 + 2(r - 1)\alpha/(4t_0)}{1 + 6\delta/4 + 2(r - 1)\alpha/(4t_0)}.$$

Since  $\alpha \leq t_0 - a_c = (1 + o(1))t_0/r$ , we have  $(r - 1)\alpha/(4t_0) \leq (r - 1)/(4r) + o(1) < 1/4$ . This together with  $\delta = o(1)$  implies that

$$\rho > \frac{6\delta/4 + 2(r - 1)\alpha/(4t_0)}{1 + 6\delta/4 + 2(r - 1)\alpha/(4t_0)} \geq \frac{\delta}{2} + \frac{(r - 1)\alpha}{2(t_0 + (r - 1)\alpha/2)}. \quad (17)$$

Now Theorem 5 implies that with probability

$$1 - \exp(-\Omega(\delta^3 p^{-1})) \stackrel{(16)}{=} 1 - \exp(-\omega(t_0)) \stackrel{\alpha \leq t_0}{=} 1 - \exp(-\omega(\alpha^2/t_0))$$

there is a component of size at least

$$\left(\rho - \frac{\delta}{4}\right) p^{-1} \stackrel{(17)}{\geq} \left(\frac{\delta}{4} + \frac{(r - 1)\alpha}{2t_0 + (r - 1)\alpha}\right) p^{-1},$$

completing the proof.  $\square$

In the following two lemmas we estimate the number of vertices with at least  $r$  neighbours in a set of size  $o(p^{-1})$  and a set of size  $\omega(p^{-1})$  separately, because estimating the probability that a vertex has at least  $r$  neighbours in such sets differs significantly.

**Lemma 11.** *Let  $B, W \subset [n]$  in  $G(n, p)$  satisfy  $|B| = (np)^{-1/(4(r-1))} p^{-1}/4$  and  $|W| = o(n)$ . With probability  $1 - \exp(-\omega(p^{-1}))$  the number of vertices in  $[n] \setminus (B \cup W)$  with at least  $r$  neighbours in  $B$  is at least  $(np)^{1/(4(r-1))} p^{-1}$ .*

*Proof.* Let  $Y_v$  be the indicator random variable that a vertex  $v \in [n] \setminus (B \cup W)$  has at least  $r$  neighbours in  $B$  and set  $Y = \sum_{v \in [n] \setminus (B \cup W)} Y_v$ . Since  $|B|p = o(1)$ , we have

$$\begin{aligned} \mathbb{P}[Y_v = 1] &= \sum_{j=r}^{|B|} \binom{|B|}{j} p^j (1 - p)^{|B|-j} = (1 + o(1)) \binom{|B|}{r} p^r (1 - p)^{|B|-r} \\ &= (1 + o(1)) \frac{|B|^r}{r!} p^r = (1 + o(1)) \frac{(np)^{-r/(4(r-1))}}{4^r r!}. \end{aligned}$$

Note that

$$1 - \frac{r}{4(r-1)} = \frac{3r-4}{4(r-1)} \stackrel{r \geq 2}{\geq} \frac{1}{2(r-1)}. \quad (18)$$

Since  $|[n] \setminus (B \cup W)| = (1 + o(1))n$ , we have

$$\mathbb{E}[Y] = (1 + o(1))n \frac{(np)^{-r/(4(r-1))}}{4^r r!} \stackrel{(18)}{\geq} (1 + o(1)) \frac{(np)^{1/(2(r-1))}}{4^r r!} p^{-1}. \quad (19)$$

Furthermore the set of random variables  $\{Y_v | v \in [n] \setminus (B \cup W)\}$  are mutually independent. Therefore, by Theorem 3 we have

$$\begin{aligned} \mathbb{P}\left[Y \leq (np)^{1/(4(r-1))} p^{-1}\right] &\leq \exp\left(-\frac{(1 + o(1))(\mathbb{E}[Y])^2}{2\mathbb{E}[Y]}\right) \\ &\stackrel{(19)}{\leq} \exp\left(-\Omega\left((np)^{1/(2(r-1))} p^{-1}\right)\right) \stackrel{np = \omega(1)}{=} \exp(-\omega(p^{-1})). \end{aligned}$$

$\square$

**Lemma 12.** Let  $C, W \subset [n]$  in  $G(n, p)$  satisfy  $|C| = (np)^{1/(4(r-1))}p^{-1}$  and  $|W| = o(n)$ . Then with probability  $1 - \exp(-\Omega(p^{-1}))$  the number of vertices in  $[n] \setminus (C \cup W)$  with at least  $r$  neighbours in  $C$  is at least  $(1 + o(1))n$ .

*Proof.* Let  $Q_v$  be the indicator random variable that a vertex  $v \in [n] \setminus (C \cup W)$  has less than  $r$  neighbours in  $C$  and set  $Q := \sum_{v \in [n] \setminus (C \cup W)} Q_v$ . Since  $|C|p = \omega(1)$ , we have

$$\mathbb{P}[Q_v = 1] = \sum_{j=0}^{r-1} \binom{|C|}{j} p^j (1-p)^{|C|-j} \leq \exp(-|C|p) \sum_{j=0}^{r-1} \frac{(|C|p)^j}{j!} \leq (1 + o(1))(|C|p)^{r-1} \exp(-|C|p).$$

Because  $|C|p = \omega(1)$ , we have

$$\exp(|C|p) = \omega((|C|p)^{5(r-1)}). \quad (20)$$

Since  $|[n] \setminus (C \cup W)| \leq n$  and  $|C|p = (np)^{1/(4(r-1))}$ , we have

$$\mathbb{E}[Q] \leq (1 + o(1))n(|C|p)^{r-1} \exp(-|C|p) \stackrel{(20)}{=} o(n(|C|p)^{-4(r-1)}) = o(n(np)^{-1}) = o(p^{-1}).$$

Note that the set of random variables  $\{Q_v | v \in [n] \setminus (C \cup W)\}$  are mutually independent and therefore, by Theorem 3 we have

$$\mathbb{P}[Q \geq p^{-1}] \leq \exp(-\Omega(p^{-1})).$$

Thus all but at most  $p^{-1}$  vertices in  $[n] \setminus (C \cup W)$  have at least  $r$  neighbours in  $C$ , which is at least  $n - |C| - |W| - p^{-1} = (1 + o(1))n$ .  $\square$

*Proof of Theorem 2.* According to Lemma 8, with probability at least

$$1 - \exp\left(- (1 + o(1)) \frac{r\alpha^2}{8(t_0 + r\alpha/3)}\right)$$

we have that  $T > t_1$  and  $|A(t_1)| \geq t_1 + (1 - 2\hat{\pi}(t_0))\alpha/4$ . Note that  $T > t_1$  implies  $|Z(t_1)| = t_1$ .

Let  $A \subset A(t_1) \setminus Z(t_1)$  with  $|A| = (1 - 2\hat{\pi}(t_0))\alpha/4 \stackrel{(12)}{=} (1 + o(1))\alpha/4$ . Lemma 9 implies that conditional on  $T > t_1$  and  $|A(t_1)| \geq t_1 + (1 - 2\hat{\pi}(t_0))\alpha/4$ , with probability at least  $1 - \exp(-\omega(\alpha^2/t_0))$  there is a set of vertices in  $[n] \setminus (Z(t_1) \cup A)$  of size at least  $\left(1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0}\right)p^{-1}$ , where every vertex in the set has at least  $r - 1$  neighbours in  $Z(t_1)$ . Select a subset  $\hat{B}$  of these vertices of size exactly

$$\left(1 + \frac{3}{4}\delta + \frac{(r-1)\alpha}{4t_0}\right)p^{-1}.$$

Note that until this point every event depends only on edges with one end in  $Z(t_1)$ .

Since the graph spanned by  $\hat{B}$  is a binomial random graph  $G((1 + 3\delta/4 + (r-1)\alpha/(4t_0))p^{-1}, p)$ , Lemma 10 implies that with probability  $1 - \exp(-\omega(\alpha^2/t_0))$  it contains a component of size at least  $(\delta/4 + (r-1)\alpha/(2t_0 + (r-1)\alpha))p^{-1}$ . Note that this event depends only on the edges with both ends in  $\hat{B}$  and therefore it is independent of the previous events. Let  $B' \subset \hat{B}$  be a component of size exactly

$$\left(\frac{\delta}{4} + \frac{(r-1)\alpha}{2t_0 + (r-1)\alpha}\right)p^{-1}.$$

Observe that if a vertex in  $A$  is connected to a vertex in  $B'$ , then every vertex in  $B'$  will become infected eventually. The probability that no vertex in  $A$  is connected to any vertex in  $B'$  is

$$(1-p)^{|A| \cdot |B'|} \leq (1-p)^{(1+o(1))|B'| \alpha/4} \leq \exp\left(- (1 + o(1)) \frac{(r-1)\alpha^2}{8t_0 + 4(r-1)\alpha}\right).$$

This event depends on edges between  $A$  and  $B'$  and thus it is independent of the previous events.

Since  $|B'| \geq \delta p^{-1}/4 \geq (np)^{-1/4(r-1)}p^{-1}/4$ , there exists a set  $B \subset B'$  with  $|B| = (np)^{-1/4(r-1)}p^{-1}/4$ . Lemma 11 implies that the number of vertices in  $[n] \setminus (Z(t_1) \cup A \cup \hat{B})$  which have at least  $r$  neighbours in  $B$  is at least  $(np)^{1/(4(r-1))}p^{-1}$  and let  $C$  be a subset of these vertices with exactly  $(np)^{1/(4(r-1))}p^{-1}$  vertices. This event depends only on edges between  $B$  and  $[n] \setminus (Z(t_1) \cup A \cup \hat{B})$  and thus it is independent of the previous events.

Finally let  $D$  be the set of vertices in  $[n] \setminus (Z(t_1) \cup A \cup B \cup C)$  which contain at least  $r$  neighbours in  $C$ . Therefore, by Lemma 12 we have that  $|D| = (1 + o(1))n$  with probability  $1 - \exp(-\Omega(p^{-1}))$ . Similarly as before this event depends only on edges which haven't been considered previously and thus it is independent of the previous events.

Since  $D \subset A_f$  and  $p^{-1} = \omega(t_0)$ , we have that the probability that almost every vertex becomes infected is at least

$$1 - \exp\left(- (1 + o(1)) \frac{r\alpha^2}{8(t_0 + r\alpha/3)}\right) - \exp\left(- (1 + o(1)) \frac{(r-1)\alpha^2}{8t_0 + 4(r-1)\alpha}\right),$$

completing the proof.  $\square$

## 6 Discussion

In a earlier weaker version of this paper [23] we consider the case when  $\delta$  is a constant, indeed  $\delta = r - 1$ . The proof for that case is simpler but the exponential tail bounds are weaker, in particular the constant in the exponential is weaker by a factor of roughly  $r - 1$ .

## References

- [1] M. Abdullah and N. Fountoulakis. A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs, 2014. arXiv:1404.4070.
- [2] M. Aizenman and J. Lebowitz. Metastability effects in bootstrap percolation. *J. Phys. A*, 21(19):3801–3813, 1988.
- [3] H. Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electron. J. Combin.*, 17(1):Research Paper 25, 20, 2010.
- [4] H. Amini. Bootstrap percolation in living neural networks. *J. Stat. Phys.*, 141(3):459–475, 2010.
- [5] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs. *Journal of Statistical Physics*, 155(1):72–92, 2014.
- [6] H. Amini, N. Fountoulakis, and K. Panagiotou. Bootstrap percolation in inhomogeneous random graphs, 2014. arXiv:1402.2815.
- [7] J. Balogh and B. Bollobás. Bootstrap percolation on the hypercube. *Probab. Theory Related Fields*, 134(4):624–648, 2006.
- [8] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364(5):2667–2701, 2012.
- [9] J. Balogh, B. Bollobás, and R. Morris. Bootstrap percolation in three dimensions. *Ann. Probab.*, 37(4):1329–1380, 2009.
- [10] J. Balogh, Y. Peres, and G. Pete. Bootstrap percolation on infinite trees and non-amenable groups. *Combin. Probab. Comput.*, 15(5):715–730, 2006.
- [11] J. Balogh and B. Pittel. Bootstrap percolation on the random regular graph. *Random Structures Algorithms*, 30(1-2):257–286, 2007.
- [12] B. Bollobás and O. Riordan. Exploring hypergraphs with martingales. *Random Structures Algorithms*, to appear. arXiv:1403.6558.
- [13] E. Candellero and N. Fountoulakis. Bootstrap percolation and the geometry of complex networks. *Stochastic Process. Appl.*, 126(1):234–264, 2016.
- [14] R. Cerf and F. Manzo. The threshold regime of finite volume bootstrap percolation. *Stochastic Process. Appl.*, 101(1):69–82, 2002.
- [15] J. Chalupa, P. L. Leath, and G. R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C: Solid State Physics*, 12:L31–L35, 1979.
- [16] H. Einarsson, J. Lengler, K. Panagiotou, F. Mousset, and A. Steger. Bootstrap Percolation with Inhibition. arXiv:1410.3291.

- [17] Chung F. and Lu L. Concentration inequalities and martingale inequalities: a survey. *Internet Math.*, 3(1):79–127, 2006.
- [18] U. Feige, M. Krivelevich, and D. Reichman. Contagious sets in random graphs, 2016. arXiv:1602.01751.
- [19] L. Fontes and R. Schonmann. Bootstrap percolation on homogeneous trees has 2 phase transitions. *J. Stat. Phys.*, 132(5):839–861, 2008.
- [20] L. Fontes, R. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Comm. Math. Phys.*, 228:495–518, 2002.
- [21] A. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab. Theory Related Fields*, 125(2):195–224, 2003.
- [22] S. Janson, T. Luczak, T. Turova, and T. Vallier. Bootstrap percolation on the random graph  $G_{n,p}$ . *Ann. Appl. Probab.*, 22(5):1989–2047, 2012.
- [23] M. Kang and T. Makai. Bootstrap percolation on  $g(n,p)$  revisited. In *Proceedings of the 27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, 2016.
- [24] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.
- [25] G.-P. Scalia-Tomba. Asymptotic final-size distribution for some chain-binomial processes. *Adv. in Appl. Probab.*, 17(3):477–495, 1985.
- [26] T. Vallier. *Random graph models and their applications*. PhD thesis, Lund Univ., 2007.
- [27] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

## 7 Appendix

Since throughout the appendix we work with martingales, we start with the well-known definition of a martingale.

**Definition 1.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space and  $\emptyset \subseteq \mathcal{F}(0) \subseteq \dots \subseteq \mathcal{F}(k) \subseteq \mathcal{F}$  be a sequence of sub- $\sigma$ -algebras. A sequence of random variables  $M_0, \dots, M_k$  forms a martingale with respect to the filtration  $\{\mathcal{F}(i)\}_{0 \leq i \leq k}$  if the following three conditions are satisfied:

- (M1) for every  $0 \leq i \leq k$ , the random variable  $M_i$  is  $\mathcal{F}(i)$ -measurable;
- (M2) for every  $0 \leq i \leq k$ ,  $\mathbb{E}[|M_i|] \leq \infty$ ;
- (M3) for every  $1 \leq i \leq k$ ,  $\mathbb{E}[M_i | \mathcal{F}(i-1)] = M_{i-1}$ .

### 7.1 Proof of Lemma 6

In order to prove Lemma 6 we let  $X(0, 1), \dots, X(n, n-a)$  and  $M(0, n-a), \dots, M(n, n-a)$  be random variables defined in Section 3. Recall that for  $(t, i) \geq (0, n-a)$  we denote the  $\sigma$ -algebra generated by  $M(0, n-a), \dots, M(t, i)$  by  $\mathcal{F}(t, i) = \sigma\left(\{M(\tau, \iota)\}_{(0, n-a) \leq (\tau, \iota) \leq (t, i)}\right)$ . Similarly we denote the  $\sigma$ -algebra generated by  $X(0, 1), \dots, X(t, i)$  by  $\mathcal{G}(t, i) = \sigma\left(\{X(\tau, \iota)\}_{(0, 1) \leq (\tau, \iota) \leq (t, i)}\right)$ .

**Lemma 13.** For any  $(0, n-a) \leq (t, i) \leq (n, n-a)$ ,  $\mathcal{G}(t, i) = \mathcal{F}(t, i)$ .

*Proof.* The proof is by induction. Since  $X(0, 1) = \dots = X(0, n-a) = 0$  and  $M(0, n-a) = 0$  we have that  $\mathcal{F}(0, n-a) = \mathcal{G}(0, n-a)$ .

Now assume that  $\mathcal{F}((t, i) - 1) = \mathcal{G}((t, i) - 1)$ . Since  $\mathcal{F}((t, i) - 1) \subseteq \mathcal{F}(t, i)$  and  $\mathcal{G}((t, i) - 1) \subseteq \mathcal{G}(t, i)$ , the result follows if we can show that  $X(t, i)$  is  $\mathcal{F}(t, i)$ -measurable and  $M(t, i)$  is  $\mathcal{G}(t, i)$ -measurable.

First we show that  $\pi(t)$  is  $\mathcal{G}((t, i) - 1)$ -measurable, and thus by our induction hypothesis it is also  $\mathcal{F}((t, i) - 1)$ -measurable. Recall that

$$|A(\tau)| \stackrel{(1)}{=} a + \sum_{i=1}^{n-a} X(\tau, i)$$

and that  $T$  is the only  $\tau$  such that  $|A(\tau)| = \tau$ . If  $\tau < t$ , then the random variables  $X(\tau, 1), \dots, X(\tau, n-a)$  are  $\mathcal{G}((t, i) - 1)$ -measurable and thus  $\mathbb{1}_{T=\tau}$  is as well. In addition the random variables  $\mathbb{1}_{T < t} = \sum_{\tau=0}^{t-1} \mathbb{1}_{T=\tau}$  and  $T \mathbb{1}_{T < t} = \sum_{\tau=0}^{t-1} \tau \mathbb{1}_{T=\tau}$  are also  $\mathcal{G}((t, i) - 1)$ -measurable.

One can express  $\pi(t) = \hat{\pi}(t)(1 - \mathbb{1}_{T < t}) + \hat{\pi}(T \mathbb{1}_{T < t}) \mathbb{1}_{T < t}$  and thus  $\pi(t)$  is also  $\mathcal{G}((t, i) - 1)$ -measurable. Recall from (4) that

$$M(t, i) - M((t, i) - 1) = \frac{X(t, i) - \pi(t)}{1 - \pi(t)} - \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)}. \quad (21)$$

The result follows from the fact that every random variable except  $M(t, i)$  in (21) is  $\mathcal{G}(t, i)$ -measurable and every variable except  $X(t, i)$  in (21) is  $\mathcal{F}(t, i)$ -measurable.  $\square$

*Proof of Lemma 6.* Clearly the conditions (M1) and (M2) are satisfied, so we only need to show (M3). Fix  $1 \leq t \leq n$  and  $1 \leq i \leq n-a$ .

First we shall show that

$$\mathbb{E} \left[ \frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1) \right] = \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)}. \quad (22)$$

To this end, observe that if  $X(t-1, i) = 1$ , then we have  $X(t, i) = 1$  and conditional on this event both sides of Equation (22) equal 1.

Now assume that  $X(t-1, i) = 0$ . When  $t > T$ , we have  $X(t, i) = X(t, i-1) = 0$  with probability 1 and by the definition of  $\pi(t)$  we have  $\pi(t) = \pi(t-1) = \hat{\pi}(T)$ . Evaluating both sides of Equation (22) gives us  $-\hat{\pi}(T)/(1 - \hat{\pi}(T))$ .

Finally we look at the case when  $t \leq T$  and  $X(t-1, i) = 0$ . Since  $t \leq T$  we have  $\pi(t) = \hat{\pi}(t)$  and thus

$$\begin{aligned} & \mathbb{E} \left[ \frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1), X(t-1, i) = 0, t \leq T \right] \\ &= \frac{\mathbb{E}[X(t, i) | \mathcal{F}((t, i) - 1), X(t-1, i) = 0, t \leq T] - \hat{\pi}(t)}{1 - \hat{\pi}(t)}. \end{aligned} \quad (23)$$

Note that  $X(t-1, i) = 0$  implies  $a + i \notin A(t-1)$  and thus  $a + i \notin Z(t)$ . When  $t \leq T$ , the random variable  $X(t, i)$  depends only on edges between  $a + i$  and  $Z(t)$  and if  $a + i \in V \setminus Z(t)$  these edges are disjoint from the edges between  $j$  and  $Z(t)$  for any  $j \neq i$ . Therefore, conditional on  $t \leq T$  we have that  $X(t, i)$  is independent of  $X(\tau, \iota)$  for every  $(0, n-a) \leq (\tau, \iota) < (t, i)$  such that  $\iota \neq i$ . This together with Lemma 13 implies

$$\mathbb{E}[X(t, i) | \mathcal{F}((t, i) - 1), X(t-1, i) = 0, t \leq T] = \mathbb{E}[X(t, i) | \sigma(\{X(\tau, i)\}_{\tau < t}), X(t-1, i) = 0, t \leq T].$$

Since  $X(t-1, i) = 0$  implies  $X(\tau, i) = 0$  for every  $\tau < t$ , we have

$$\mathbb{E}[X(t, i) | \sigma(\{X(\tau, i)\}_{\tau < t}), X(t-1, i) = 0, t \leq T] = \mathbb{E}[X(t, i) | X(t-1, i) = 0, t \leq T].$$

Note that

$$\mathbb{P}[X(t, i) = 0 | X(t-1, i) = 0, t \leq T] = \frac{1 - \hat{\pi}(t)}{1 - \hat{\pi}(t-1)} = 1 - \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)}.$$

Therefore we obtain

$$\mathbb{E}[X(t, i) | X(t-1, i) = 0, t \leq T] = \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)}, \quad (24)$$

which together with (23) gives

$$\begin{aligned} \mathbb{E} \left[ \frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1), X(t-1, i) = 0, t \leq T \right] &= -\frac{\hat{\pi}(t)}{1 - \hat{\pi}(t)} \cdot \frac{1 - \hat{\pi}(t)}{1 - \hat{\pi}(t-1)} + 1 \cdot \frac{\hat{\pi}(t) - \hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)} \\ &= -\frac{\hat{\pi}(t-1)}{1 - \hat{\pi}(t-1)}, \end{aligned}$$

implying (22). Finally using (21) and (22) we have

$$\begin{aligned} & \mathbb{E}[M(t, i) - M((t, i) - 1) | \mathcal{F}((t, i) - 1)] \\ & \stackrel{(21)}{=} \mathbb{E} \left[ \frac{X(t, i) - \pi(t)}{1 - \pi(t)} \middle| \mathcal{F}((t, i) - 1) \right] - \frac{X(t-1, i) - \pi(t-1)}{1 - \pi(t-1)} \stackrel{(22)}{=} 0. \end{aligned}$$

□

## 7.2 Proof of Theorem 4

We need the following well-known results from conditional expectation [27]. Let  $X, Y$  be random variables with  $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$  and let  $\mathcal{G}$  be a sub- $\sigma$ -algebra of  $\mathcal{F}$ . Then the following holds.

(C1)  $\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X]$ ;

(C2)  $\mathbb{E}[aX + bY | \mathcal{G}] = a\mathbb{E}[X | \mathcal{G}] + b\mathbb{E}[Y | \mathcal{G}]$  for  $a, b \in \mathbb{R}$ ;

(C3) If  $Y$  is  $\mathcal{G}$ -measurable, then  $\mathbb{E}[Y | \mathcal{G}] = Y$ ;

(C4) If  $X, Y > 0$ ,  $\mathbb{E}[XY] < \infty$ , and  $Y$  is  $\mathcal{G}$ -measurable, then  $\mathbb{E}[XY | \mathcal{G}] = Y\mathbb{E}[X | \mathcal{G}]$ .

Note that if  $\mathcal{G} = \emptyset$ , then  $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$ . We also need the following result of McDiarmid [24].

**Lemma 14.** *Let  $g(x) = 2 \sum_{\ell=2}^{\infty} x^{\ell-2} / \ell!$ . Then  $g(x)$  is monotone increasing for  $x \in \mathbb{R}_+$  and in addition  $g(x) < (1 - x/3)^{-1}$  for  $x < 3$ .*



*Proof of Theorem 4.* Note that for every  $1 \leq i \leq k$ , if  $\sigma_i^2 = 0$ , then  $\text{Var}[M_i|\mathcal{F}(i-1)] = 0$ . Therefore, if  $\sigma_i^2 = 0$  for every  $1 \leq i \leq k$ , then we have  $M_k = M_{k-1} = M_0 = m_0$ . So for the remainder of the proof we may assume that  $\sum_{i=1}^k \sigma_i^2 > 0$ .

For  $1 \leq i \leq k$  and  $t > 0$  we have

$$\mathbb{E} \left[ e^{t(M_i - M_{i-1})} | \mathcal{F}(i-1) \right] = \mathbb{E} \left[ \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} (M_i - M_{i-1})^\ell | \mathcal{F}(i-1) \right]. \quad (25)$$

Note that

$$\mathbb{E}[M_i - M_{i-1} | \mathcal{F}(i-1)] \stackrel{(C2)}{=} \mathbb{E}[M_i | \mathcal{F}(i-1)] - \mathbb{E}[M_{i-1} | \mathcal{F}(i-1)] \stackrel{(C3), (M1)}{=} \mathbb{E}[M_i | \mathcal{F}(i-1)] - M_{i-1} \stackrel{(M3)}{=} 0.$$

This together with (25) and (C2) implies

$$\begin{aligned} \mathbb{E} \left[ e^{t(M_i - M_{i-1})} | \mathcal{F}(i-1) \right] &= 1 + \mathbb{E} \left[ \sum_{\ell=2}^{\infty} \frac{t^\ell}{\ell!} (M_i - M_{i-1})^\ell | \mathcal{F}(i-1) \right] \\ &= 1 + \mathbb{E} \left[ \frac{t^2}{2} (M_i - M_{i-1})^2 g(t(M_i - M_{i-1})) | \mathcal{F}(i-1) \right]. \end{aligned}$$

By the condition (i) of Theorem 4 we have  $M_i - M_{i-1} \leq m$  and by Lemma 14 the function  $g$  is monotone increasing and thus  $g(t(M_i - M_{i-1})) \leq g(tm)$ , leading to

$$\begin{aligned} \mathbb{E} \left[ e^{t(M_i - M_{i-1})} | \mathcal{F}(i-1) \right] &\leq 1 + \mathbb{E} \left[ \frac{t^2}{2} (M_i - M_{i-1})^2 g(tm) | \mathcal{F}(i-1) \right] \\ &\stackrel{(C2)}{=} 1 + \frac{t^2}{2} g(tm) \mathbb{E} [(M_i - M_{i-1})^2 | \mathcal{F}(i-1)] \\ &\stackrel{(ii)}{\leq} 1 + \frac{t^2}{2} g(tm) \sigma_i^2 \\ &\leq \exp \left( \frac{t^2}{2} g(tm) \sigma_i^2 \right). \end{aligned} \quad (26)$$

Clearly

$$\mathbb{E} \left[ e^{tM_i} | \mathcal{F}(i-1) \right] = \mathbb{E} \left[ e^{t(M_i - M_{i-1})} e^{tM_{i-1}} | \mathcal{F}(i-1) \right].$$

Note that both  $e^{t(M_i - M_{i-1})} > 0$  and  $e^{tM_{i-1}} > 0$ . In addition (i) implies  $M_i \leq m_0 + im$  for any  $0 \leq i \leq k$ . Therefore  $e^{tM_i}$  is also bounded from above and thus  $\mathbb{E}[e^{tM_i}] < \infty$ . By (M1) and (C4) we have

$$\mathbb{E} \left[ e^{tM_i} | \mathcal{F}(i-1) \right] = e^{tM_{i-1}} \mathbb{E} \left[ e^{t(M_i - M_{i-1})} | \mathcal{F}(i-1) \right] \stackrel{(26)}{\leq} e^{tM_{i-1}} \exp \left( \frac{t^2}{2} g(tm) \sigma_i^2 \right). \quad (27)$$

Now we will show by induction that for each  $0 \leq i \leq k$

$$\mathbb{E} \left[ e^{tM_i} \right] \leq \exp \left( tm_0 + \frac{t^2}{2} g(tm) \sum_{j=1}^i \sigma_j^2 \right).$$

Observe that this holds for  $i = 0$ . Assume that it holds for  $i - 1$ :

$$\mathbb{E}[e^{tM_{i-1}}] \leq \exp \left( tm_0 + \frac{t^2}{2} g(tm) \sum_{j=1}^{i-1} \sigma_j^2 \right). \quad (28)$$

Then we have

$$\begin{aligned} \mathbb{E} \left[ e^{tM_i} \right] &\stackrel{(C1)}{=} \mathbb{E} \left[ \mathbb{E} \left[ e^{tM_i} | \mathcal{F}(i-1) \right] \right] \\ &\stackrel{(27)}{\leq} \mathbb{E} \left[ e^{tM_{i-1}} \exp \left( \frac{t^2}{2} g(tm) \sigma_i^2 \right) \right] \\ &\stackrel{(C2)}{=} \exp \left( \frac{t^2}{2} g(tm) \sigma_i^2 \right) \mathbb{E}[e^{tM_{i-1}}] \\ &\stackrel{(28)}{\leq} \exp \left( tm_0 + \frac{t^2}{2} g(tm) \sum_{j=1}^i \sigma_j^2 \right). \end{aligned}$$

In particular, we have

$$\mathbb{E}[e^{t(M_k - m_0)}] \leq \exp\left(\frac{t^2}{2}g(tm)\sum_{i=1}^k\sigma_i^2\right). \quad (29)$$

By Markov's inequality we have that for any  $\lambda, t > 0$

$$\begin{aligned} \mathbb{P}[M_k - m_0 \geq \lambda] &= \mathbb{P}[e^{t(M_k - m_0)} \geq e^{t\lambda}] \\ &\leq e^{-t\lambda}\mathbb{E}[e^{t(M_k - m_0)}] \\ &\stackrel{(29)}{\leq} \exp\left(-t\lambda + \frac{t^2}{2}g(tm)\sum_{i=1}^k\sigma_i^2\right). \end{aligned} \quad (30)$$

Set

$$t = \frac{\lambda}{\sum_{i=1}^k\sigma_i^2 + m\lambda/3}. \quad (31)$$

Recall that  $\sum_{i=1}^k\sigma_i^2 > 0$  and thus

$$tm = \frac{m\lambda}{\sum_{i=1}^k\sigma_i^2 + m\lambda/3} < \frac{m\lambda}{m\lambda/3} = 3.$$

Therefore by Lemma 14 we have  $g(tm) \leq (1 - tm/3)^{-1}$  and

$$\begin{aligned} \mathbb{P}[M_k - m_0 \geq \lambda] &\stackrel{(30)}{\leq} \exp\left(\frac{t^2}{2}\frac{1}{1 - tm/3}\sum_{j=1}^k\sigma_j^2 - t\lambda\right) \\ &\stackrel{(31)}{=} \exp\left(-\frac{\lambda^2}{2\left(\sum_{i=1}^k\sigma_i^2 + m\lambda/3\right)}\right). \end{aligned}$$

□